



**Original citation:**

Botta, Federico, Moat, Helen Susannah, Stanley, H. Eugene and Preis, Tobias. (2015) Quantifying stock return distributions in financial markets. PLoS One, 10 (9). e0135600.

**Permanent WRAP url:**

<http://wrap.warwick.ac.uk/74548>

**Copyright and reuse:**


The Warwick Research Archive Portal (WRAP) makes this work of researchers of the University of Warwick available open access under the following conditions.

This article is made available under the Creative Commons Attribution 4.0 International license (CC BY 4.0) and may be reused according to the conditions of the license. For more details see: <http://creativecommons.org/licenses/by/4.0/>

**A note on versions:**

The version presented in WRAP is the published version, or, version of record, and may be cited as it appears here.

For more information, please contact the WRAP Team at: [publications@warwick.ac.uk](mailto:publications@warwick.ac.uk)

warwick**publications**wrap  
  
highlight your research

<http://wrap.warwick.ac.uk>

RESEARCH ARTICLE

# Quantifying Stock Return Distributions in Financial Markets

Federico Botta<sup>1,2\*</sup>, Helen Susannah Moat<sup>2</sup>, H. Eugene Stanley<sup>3</sup>, Tobias Preis<sup>2</sup>

**1** Centre for Complexity Science, University of Warwick, Coventry, CV4 7AL, United Kingdom, **2** Data Science Lab, Behavioural Science, Warwick Business School, University of Warwick, Coventry, CV4 7AL, United Kingdom, **3** Center for Polymer Studies and Department of Physics, Boston University, Boston, MA 02215, United States of America

\* [f.botta@warwick.ac.uk](mailto:f.botta@warwick.ac.uk)



## OPEN ACCESS

**Citation:** Botta F, Moat HS, Stanley HE, Preis T (2015) Quantifying Stock Return Distributions in Financial Markets. PLoS ONE 10(9): e0135600. doi:10.1371/journal.pone.0135600

**Editor:** Yanguang Chen, Peking University, CHINA

**Received:** May 19, 2015

**Accepted:** July 23, 2015

**Published:** September 1, 2015

**Copyright:** © 2015 Botta et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** Relevant data were obtained by the authors from the third party Wharton Research Data Services. Raw data sets from the Trades and Quotes database are available from the following URL: <https://wrds-web.wharton.upenn.edu/wrds/>.

**Funding:** FB acknowledges the support of UK EPSRC EP/E501311/1. HSM and TP acknowledge the support of the Research Councils UK Grant EP/K039830/1. HES acknowledges the support of IARPA and NSF (grants 1411158 and 1452061).

**Competing Interests:** The authors have declared that no competing interests exist.

## Abstract

Being able to quantify the probability of large price changes in stock markets is of crucial importance in understanding financial crises that affect the lives of people worldwide. Large changes in stock market prices can arise abruptly, within a matter of minutes, or develop across much longer time scales. Here, we analyze a dataset comprising the stocks forming the Dow Jones Industrial Average at a second by second resolution in the period from January 2008 to July 2010 in order to quantify the distribution of changes in market prices at a range of time scales. We find that the tails of the distributions of logarithmic price changes, or returns, exhibit power law decays for time scales ranging from 300 seconds to 3600 seconds. For larger time scales, we find that the distributions tails exhibit exponential decay. Our findings may inform the development of models of market behavior across varying time scales.

## Introduction

Complex movements in stock market prices affect the personal fortunes of people around the globe [1–5]. An ability to more accurately quantify and predict such changes would allow us to gain more insights into how financial crises arise [6] and provide greater empirical basis for the development of theories of financial market behavior [7–13].

The financial markets were however one of the earliest sources of large scale datasets on human behaviour, where such data have recently become the focus of the new field of computational social science [14–24]. A vast amount of data on financial decisions made in stock markets is therefore available [25–29]. Previous studies have shown that distributions of returns observed in empirical data are consistent with power law decay [30–42], in contrast with widely used models that assume Gaussian behavior of these returns. Power law behavior has also been observed in other economical and financial sectors of society [43, 44].

Changes in stock market prices can occur at a range of different time scales. Here, we analyze a large dataset of stocks forming the Dow Jones Industrial Average (DJIA) at a second-by-second resolution for a range of different time scales in order to quantify the distribution of returns. We provide evidence that while the distribution of returns exhibits power law behavior

at small time scale, exponential behavior is observed at larger time scales. We find analogous results when restricting our analysis to volatile trading periods. Our findings could help to gain insight into changes in stock market prices in shorter periods and longer periods and provide further empirical basis for the development of new models of market behavior.

## Results

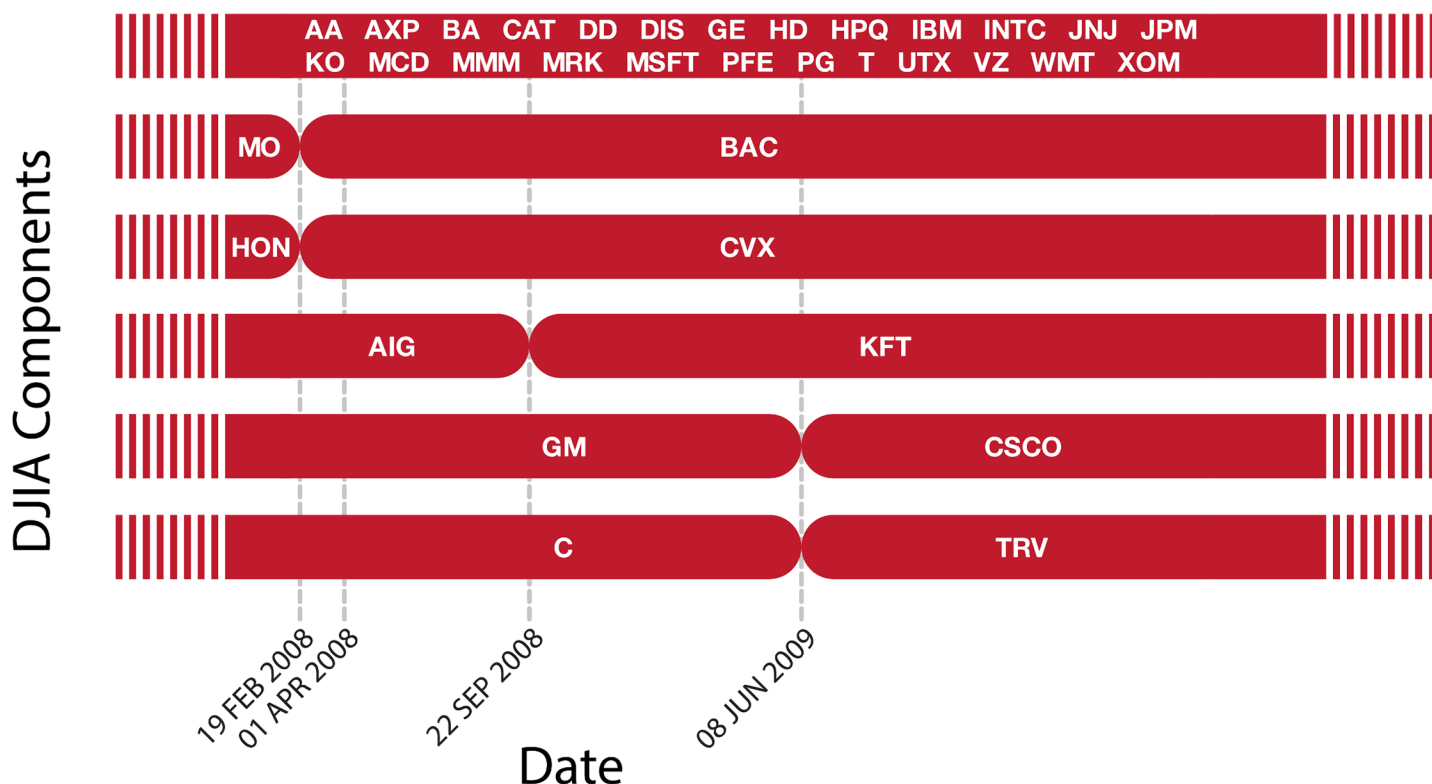
The DJIA is a U.S. benchmark index that consists of 30 different stocks. For all 30 stocks, we retrieve price time series with a second by second resolution from the Trade and Quote (TAQ) database provided by Wharton Research Data Services (WRDS). Our dataset covers the period from 2 January 2008 to 30 July 2010 comprising a total of 647 trading days. Fig 1 shows the various components of the DJIA. As five stocks were replaced during this period, we focus on the 25 components that were consistently part of the DJIA between 02 January 2008 and 30 July 2010.

We define returns as the relative logarithmic change in price of a given stock  $i$  at a given time  $t$ :

$$r_i(t) = \log(p_i(t + \Delta t)) - \log(p_i(t)) \quad i = 1, \dots, 25$$

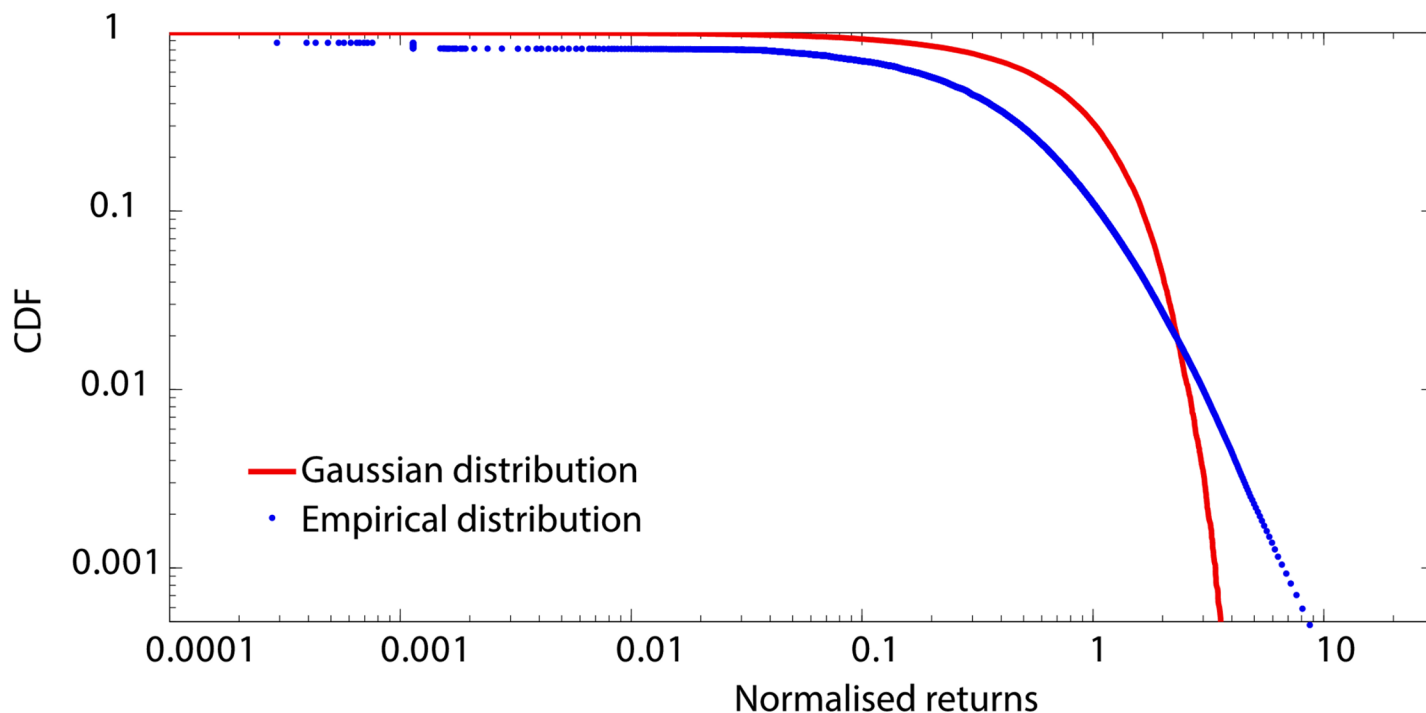
where  $\Delta t$  is the time lag between price observations. As a trading day starts at 9:30 and ends at 16:00 local time,  $\Delta t$  is constrained to be at most 6 hours and 30 minutes.

We compute the standardized distribution of the returns for the 25 components of the DJIA that we consider. We conduct separate analyses of the cumulative distribution function (CDF)



**Fig 1. Components of the DJIA.** Here we depict the components of the DJIA in the time period between 02 January 2008 to 30 July 2010. Dashed vertical lines correspond to changes in the stocks forming the DJIA. In our analysis, we focus on the 25 stocks that were part of the DJIA during the period of analysis. Stocks are labelled using ticker symbols that uniquely identify the company name, as used by the stock exchange.

doi:10.1371/journal.pone.0135600.g001



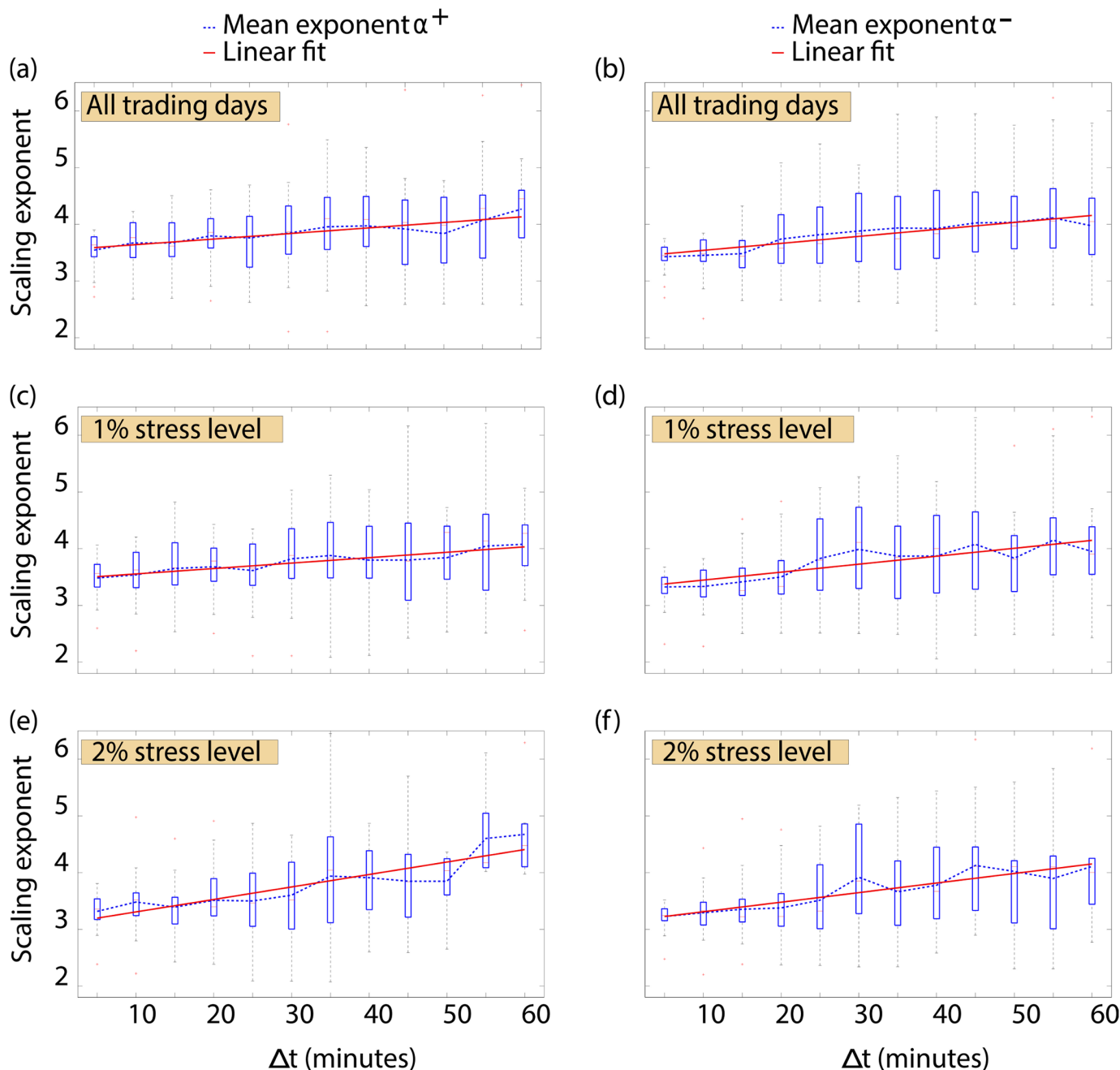
**Fig 2. Empirical distribution of normalised returns for *American Express*.** We build returns distributions for the 25 stocks of the DJIA for different time lags across the full period of analysis. We standardize each distribution by subtracting the mean return from each observation and dividing by the standard deviation. We depict in blue the cumulative distribution function of the positive component of the return distributions for *American Express* for a time lag of 300 seconds. We depict in red the positive tail of a Gaussian distribution with mean zero and standard deviation one. We observe a strong deviation of the empirical distribution from the Gaussian distribution. Instead, visual inspection of the distribution tail reveals consistency with a linear relationship on a log-log scale. This provides initial evidence for possible power law behavior at this time scale.

doi:10.1371/journal.pone.0135600.g002

of the positive and negative component of the distribution of returns. Fig 2 depicts the positive CDF for *American Express* for  $\Delta t = 300$  seconds and compares this to a Gaussian distribution. Note that the empirical distribution strongly deviates from the Gaussian distribution and provides initial evidence for power law behavior. We perform a statistical analysis to check the consistency of the tails of the empirical distributions with power law behavior across different time scales, as proposed by Clauset, Shalizi and Newman [45] and detailed in the Methods section.

### Changes in power law behavior as $\Delta t$ increases

A power law probability distribution is a probability distribution in which the probability of an event decays as a negative power of the event. The distribution function is characterized by a scaling exponent. Distributions of returns typically exhibit power law decay in the tail of the distribution. Here, we want to understand how the exact nature of power law behavior depends on the time lag between price observations. We analyze all 25 stock price time series and use a time lag  $\Delta t$  ranging from 300 to 3,600 seconds. We investigate how the scaling exponent changes as a function of the time lag between price observations. We depict the exponent for the tails of the positive (denoted as  $\alpha^+$ ; Fig 3a) and negative (denoted as  $\alpha^-$ ; Fig 3b) returns distributions obtained when analyzing data from all trading days. For both positive and negative tails, we find that the mean scaling exponent increases with the time lag  $\Delta t$  ( $\alpha^+$ : Adjusted  $R^2 = 0.802$ ,  $N = 12$ ,  $p < 0.001$ , ordinary least squares regression;  $\alpha^-$ : Adjusted  $R^2 = 0.839$ ,  $N = 12$ ,



**Fig 3. Relationship between  $\Delta t$  and the scaling exponent for the empirical tails of return distributions.** (a) We investigate the relationship between the time lag between price observations used to build the returns distribution and the scaling exponents of the tails of distributions. We consider here the tails of the positive component of the distributions obtained when analyzing all trading days present in our dataset. We find that the mean scaling exponent increases as  $\Delta t$  increases (Adjusted  $R^2 = 0.802$ ,  $N = 12$ ,  $p < 0.001$ , ordinary least squares regression) (b) In a similar fashion, we observe that when analyzing all trading days the mean scaling exponent for the tail of the negative component of the distributions increases with the time lag (Adjusted  $R^2 = 0.839$ ,  $N = 12$ ,  $p < 0.001$ , ordinary least squares regression) (c) We now restrict our analysis to trading days on which the prices of stocks have changed by more than 1%. We find that the mean scaling exponent of positive tails consistent with power law behavior increases with  $\Delta t$  (Adjusted  $R^2 = 0.856$ ,  $N = 12$ ,  $p < 0.001$ , ordinary least squares regression) (d) Under 1% stress, an increase in the time lag  $\Delta t$  results again in an increase of the mean scaling exponent for the tails of the negative returns distributions (Adjusted  $R^2 = 0.729$ ,  $N = 12$ ,  $p < 0.001$ , ordinary least squares regression) (e) We now perform the same analysis for days on which the prices of stocks have changed by more than 2%. The mean scaling exponent for the tails of the positive component of the distributions again shows an increase with increasing  $\Delta t$  (Adjusted  $R^2 = 0.782$ ,  $N = 12$ ,  $p < 0.001$ , ordinary least squares regression) (f) Similarly, the mean scaling exponent for the tails of negative returns distributions at the 2% stress level increases as the time lag  $\Delta t$  between price observations increases (Adjusted  $R^2 = 0.836$ ,  $N = 12$ ,  $p < 0.001$ , ordinary least squares regression).

doi:10.1371/journal.pone.0135600.g003

$p < 0.001$ , ordinary least squares regression):

$$\alpha^+ = 0.010(\pm 0.001)\Delta t + 3.54(\pm 0.05)$$

$$\alpha^- = 0.012(\pm 0.001)\Delta t + 3.42(\pm 0.06)$$

We find similar slopes for the positive and negative tails, which suggests that both exponents  $\alpha^+$  and  $\alpha^-$  vary in a similar fashion as a function of the time lag  $\Delta t$ . Our results suggest that the probability of finding large price changes is underestimated by a Gaussian distribution and better quantified by a power law distribution, in line with a range of findings reported in the field of econophysics [30–42]. Previous findings for US markets have highlighted that stock returns may follow an inverse cubic law [31]. The analysis of different stock markets, such as the Warsaw Stock Exchange in Poland or the Australian Stock Exchange, has uncovered different power law regimes deviating from the inverse cubic law [38, 39]. By selecting appropriate cutoff values in the distributions under analysis, stocks from the Mexican Stock Market index exhibit a power law decay close to an inverse cubic law [40]. Analogous results have also been observed when analysing daily returns in Chinese stock markets [46, 47].

It remains unclear, however, whether these findings hold for subsets of the price time series in which extreme price movements occur. We therefore restrict our analysis to price observations recorded on trading days on which the corresponding stock gained or lost more than 1% on a daily basis. We refer to this as a stress level of 1%. Fig 3c and 3d depict the relationship between the power law exponents and the time lag  $\Delta t$  between price observations on trading days on which the market experienced a stress level of at least 1%.

We again find that an increase in  $\Delta t$  results in an increase of the mean scaling exponent ( $\alpha^+$ : Adjusted  $R^2 = 0.856$ ,  $N = 12$ ,  $p < 0.001$ , ordinary least squares regression;  $\alpha^-$ : Adjusted  $R^2 = 0.729$ ,  $N = 12$ ,  $p < 0.001$ , ordinary least squares regression):

$$\alpha^+ = 0.010(\pm 0.001)\Delta t + 3.46(\pm 0.04)$$

$$\alpha^- = 0.014(\pm 0.003)\Delta t + 3.31(\pm 0.09)$$

We notice a strong similarity between the relationship between the scaling exponent and  $\Delta t$  in this scenario and in the scenario where we consider all trading days. In a parallel analysis, we consider a 2% stress level (Fig 3e and 3f). We find that the mean scaling exponent increases with the time lag  $\Delta t$  between price observations ( $\alpha^+$ : Adjusted  $R^2 = 0.782$ ,  $N = 12$ ,  $p < 0.001$ , ordinary least squares regression;  $\alpha^-$ : Adjusted  $R^2 = 0.836$ ,  $N = 12$ ,  $p < 0.001$ , ordinary least squares regression):

$$\alpha^+ = 0.022(\pm 0.003)\Delta t + 3.09(\pm 0.13)$$

$$\alpha^- = 0.017(\pm 0.002)\Delta t + 3.14(\pm 0.08)$$

At a stress level of 3%, we again observe that the scaling exponent increases as we increase the time lag  $\Delta t$  ( $\alpha^+$ : Adjusted  $R^2 = 0.573$ ,  $N = 12$ ,  $p < 0.05$ , ordinary least squares regression;  $\alpha^-$ : Adjusted  $R^2 = 0.458$ ,  $N = 12$ ,  $p < 0.05$ , ordinary least squares regression):

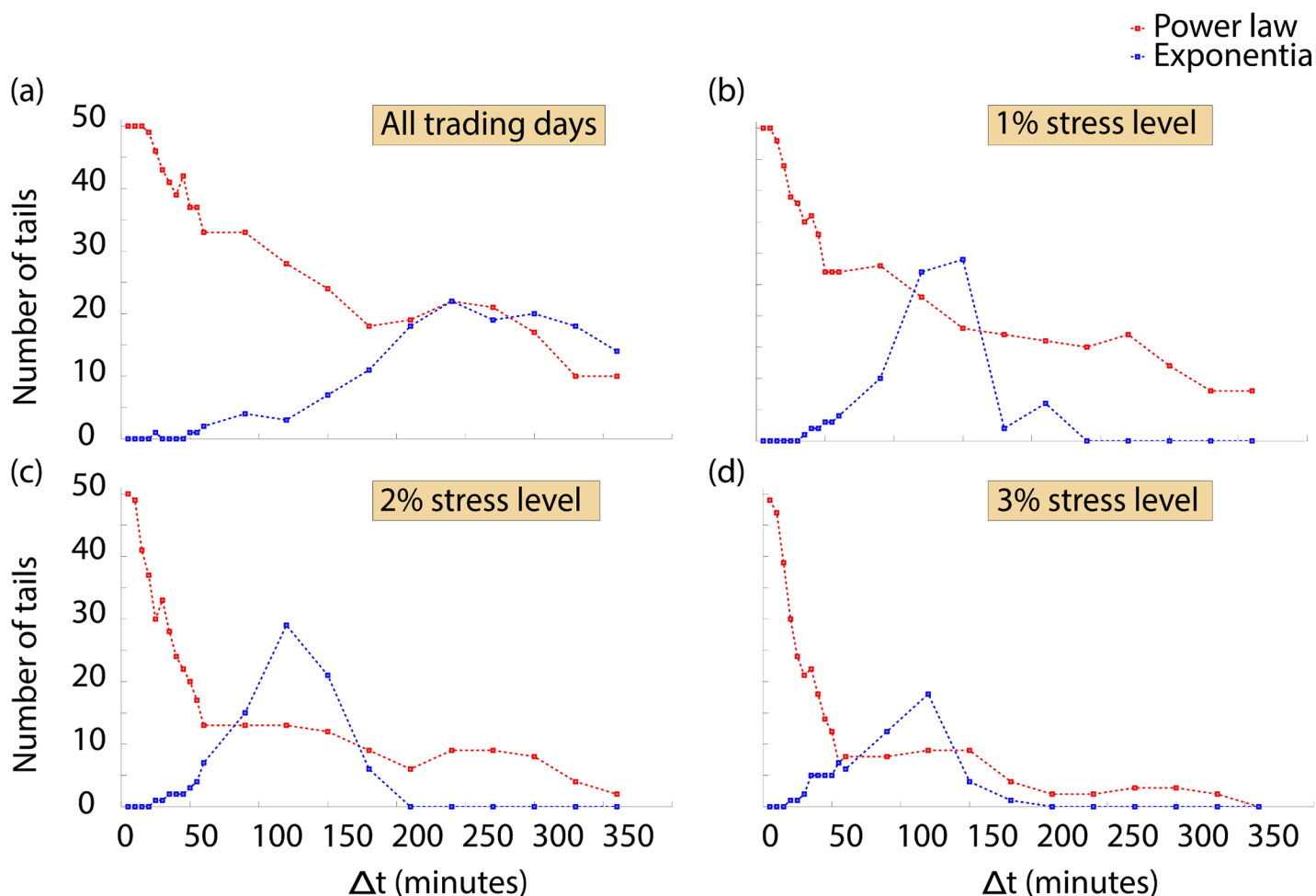
$$\alpha^+ = 0.066(\pm 0.017)\Delta t + 2.04(\pm 0.61)$$

$$\alpha^- = 0.016(\pm 0.005)\Delta t + 2.91(\pm 0.19)$$

## Evidence of exponential decay at larger values of $\Delta t$

We observe that for  $\Delta t > 60$  minutes the number of tails consistent with power law behavior decreases (Fig 4a). We investigate this change in behavior at a range of time scales and analyze whether we start to observe consistency with exponential decay. Exponential decay has already been observed in daily returns of stocks from the National Stock Exchange in the Indian stock market [48].

Fig 4a depicts how the number of distributions consistent with either power law behavior or exponential decay varies with  $\Delta t$ . At small time scales, the tail of most distributions is consistent with power law behavior. As we increase the time lag between price observations, we



**Fig 4. Consistency of empirical returns distributions with power law and exponential decay.** (a) For  $\Delta t > 60$  minutes, we note a decrease in the number of tails consistent with power law decay. We investigate whether the tails of the returns distributions are consistent with power law behavior or exponential decay using the Kolmogorov-Smirnov statistic, as described in the methods section. We first consider all trading days present in our dataset. At short time scales, we observe that the tails of most empirical distributions are consistent with power law behavior. As we increase the time lag, the number of tails consistent with power law behavior decreases and we see an increase in the number of tails of returns distributions that are consistent with exponential decay. We depict here the overall number of tails, both for the positive and negative returns distributions, for the 25 components of the DJIA. (b) We consider transaction days on which the prices of stocks have changed by more than 1%. We refer to this as a stress level of 1%. In this scenario, the number of tails consistent with power law behavior decreases more sharply. Consistency with exponential decay appears when  $\Delta t$  is roughly 2 hours. (c) In a similar fashion, when we consider a stress level of 2%, we again observe a sharp decrease in the number of distributions consistent with power law behavior. We also find an increase in the number of tails consistent with exponential decay again when  $\Delta t$  is roughly 2 hours. (d) Under a stress level of 3%, the number of empirical distributions consistent with power law behavior decreases more quickly than in the other scenarios. The number of tails consistent with exponential decay peaks at a lower number than in other scenarios, but is again highest when  $\Delta t$  is roughly two hours, similar to other scenarios.

doi:10.1371/journal.pone.0135600.g004



observe an increase in the number of tails consistent with exponential decay. At the 1% stress level, the decrease in the number of tails consistent with power law is sharper and we find a peak in the number of tails consistent with exponential decay when  $\Delta t$  is roughly 2 hours (Fig 4b). As we increase the stress level, we find that the number of tails consistent with power law behavior decreases even more sharply. The number of tails consistent with exponential decay exhibits a peak at similar time scales, but peaks at a lower number than observed at the 1% stress level (Fig 4c and 4d).

## Conclusions

Large changes in stock market prices can occur at a range of time scales, arising within minutes or developing across longer time scales. Our findings provide evidence that in different scenarios the scaling exponent of those distributions consistent with power law behavior increases with the time lag between price observations. As this time lag increases, we observe that the number of return distributions consistent with power law behavior decreases sharply. At a time lag of roughly two hours, we also find an increase in the number of distributions which are consistent with exponential decay. Our results are consistent with the hypothesis that changes in stock market prices have different behaviors at different time scales. We observe that these results hold in different scenarios of the market, both when we consider all trading days, but also when restricting our analysis to scenarios with different stress levels. We suggest that our analysis may provide further empirical insights for the development of models of market behavior.

## Methods

To check the consistency of the tails of observed empirical distributions with power law behavior, we follow the procedure proposed in Ref. [45], which we summarise again here.

A power law is a distribution of the form:

$$p(x) = \frac{\alpha - 1}{x_{\min}} \left( \frac{x}{x_{\min}} \right)^{-\alpha}$$

where  $\alpha$  is the scaling exponent. We require  $\alpha > 1$  for this to be a Probability Distribution Function (PDF).  $x_{\min}$  is the lower bound of the power law behavior. We estimate the scaling exponent  $\alpha$  using the maximum likelihood estimator (MLE). Assuming we have  $n$  observations of  $x_i (i = 1, \dots, n)$  which are independent and identically distributed random variables, the likelihood function, which represents the probability of observing the data given the parameter, is given by:

$$p(x|\alpha) = \prod_{i=1}^n \frac{\alpha - 1}{x_{\min}} \left( \frac{x_i}{x_{\min}} \right)^{-\alpha}$$

We then maximise this probability to find the MLE estimator for the scaling exponent:

$$\hat{\alpha} = 1 + n \left[ \sum_{i=1}^n \ln \frac{x_i}{x_{\min}} \right]$$

We measure distances between distribution using the Kolmogorov-Smirnov statistic (KS statistic):

$$D = \max_{x \geq x_{\min}} |E(x) - F(x)|$$



where  $E(x)$  is the empirical CDF and  $F(x)$  is the best fit of the data. We determine the lower bound  $x_{\min}$  by choosing the value that minimizes the distance between the empirical distribution and the fitted distribution as measured by the KS statistic.

Once we have determined the lower bound  $x_{\min}$  and the scaling exponent  $\alpha$ , we then check the consistency of the hypothesis of power law behavior in the observed empirical distributions. We construct the empirical tails choosing a bin size such that we have 1,000 data points in each tail. We then compare the KS statistic observed for the empirical data when compared to a fitted power law distribution with the KS statistic obtained for the synthetic data when compared to a fitted power law distribution. We obtain a  $p$ -value by counting the number of times that the synthetic KS statistic is larger than the empirical KS statistic. We generate 1,000 synthetic data sets and make the conservative choice of accepting our hypothesis of consistency with power law behavior if the  $p$ -value is larger than 0.1.

To determine whether the distribution is consistent with exponential decay, we perform a parallel analysis fitting the data to an exponential distribution instead of a power law probability distribution. We then generate synthetic data from the fitted distribution in the same manner as previously described. We evaluate whether our data are consistent with exponential decay by comparing the empirical data to the synthetic data using KS statistics as described above.

## Author Contributions

Conceived and designed the experiments: FB HSM HES TP. Analyzed the data: FB HSM HES TP. Wrote the paper: FB HSM HES TP.

## References

1. Sornette D. Why stock markets crash: critical events in complex financial systems. Princeton, NJ: Princeton University Press; 2004.
2. Farmer JD, Joshi S. The price dynamics of common trading strategies. *J Econ Behav Organ.* 2002; 49: 149–171 doi: [10.1016/S0167-2681\(02\)00065-3](https://doi.org/10.1016/S0167-2681(02)00065-3)
3. Voit J. The statistical mechanics of financial markets. Heidelberg: Springer; 2005.
4. Paul W, Baschnagel J. Stochastic processes: from physics to finance. Switzerland: Springer International Publishing; 2013.
5. Abergel F, Chakrabarti BK, Charabarti A, Mitra M. Econophysics of order-driven markets. Milan: Springer; 2011.
6. Lux T, Westerhoff F. Economics crisis. *Nat Phys.* 2009; 5: 2–3 doi: [10.1038/nphys1163](https://doi.org/10.1038/nphys1163)
7. Farmer JD, Foley D. The economy needs agent-based modeling. *Nature.* 2009; 460: 685–686 doi: [10.1038/460685a](https://doi.org/10.1038/460685a) PMID: [19661896](https://pubmed.ncbi.nlm.nih.gov/19661896/).
8. Feng L, Baowen L, Podobnik B, Preis T, Stanley HE. Linking agent-based models and stochastic models of financial markets. *Proc Natl Acad Sci USA.* 2012; 109: 8388–8393 doi: [10.1073/pnas.1205013109](https://doi.org/10.1073/pnas.1205013109) PMID: [22586086](https://pubmed.ncbi.nlm.nih.gov/22586086/).
9. Petersen AM, Wang F, Havlin S, Stanley HE. Market dynamics immediately before and after financial shocks: quantifying the Omori, productivity, and Bath laws. *Phys Rev E.* 2010; 82: 036114 doi: [10.1103/PhysRevE.82.036114](https://doi.org/10.1103/PhysRevE.82.036114)
10. Hommes CH. Modeling the stylized facts in finance through simple nonlinear adaptive systems. *Proc Natl Acad Sci USA.* 2002; 99: 7221–7228 doi: [10.1073/pnas.082080399](https://doi.org/10.1073/pnas.082080399) PMID: [12011401](https://pubmed.ncbi.nlm.nih.gov/12011401/).
11. Preis T, Golke S, Paul W, Schneider JJ. Multi-agent-based order book model of financial markets. *Europhys Lett.* 2006; 75: 510 doi: [10.1209/epl/i2006-10139-0](https://doi.org/10.1209/epl/i2006-10139-0)
12. Mantegna R, Stanley HE. An Introduction to Econophysics: Correlations and Complexity in Finance. Cambridge, MA: Cambridge University Press; 2000.
13. Carbone A, Castelli G, Stanley HE. Time-dependent Hurst exponent in financial time series. *Physica A.* 2004; 344: 267–271 doi: [10.1016/j.physa.2004.06.130](https://doi.org/10.1016/j.physa.2004.06.130)
14. Lazer D et al. Computational social science. *Science.* 2009; 323: 721–723 doi: [10.1126/science.1167742](https://doi.org/10.1126/science.1167742) PMID: [19197046](https://pubmed.ncbi.nlm.nih.gov/19197046/).

15. King G. Ensuring the data rich future of the social sciences. *Science*. 2011; 331: 719–721 doi: [10.1126/science.1197872](https://doi.org/10.1126/science.1197872) PMID: [21311013](https://pubmed.ncbi.nlm.nih.gov/21311013/).
16. de Montjoye Y-A, Radaelli L, Singh VK, Pentland A. Unique in the shopping mall: On the reidentifiability of credit card metadata. *Science*. 2015; 347: 536–539 doi: [10.1126/science.1256297](https://doi.org/10.1126/science.1256297) PMID: [25635097](https://pubmed.ncbi.nlm.nih.gov/25635097/).
17. Moat HS, Preis T, Olivola CY, Liu C, Chater N. Using big data to predict collective behavior in the real world. *Behav Brain Sci*. 2014; 37: 92–93 doi: [10.1017/S0140525X13001817](https://doi.org/10.1017/S0140525X13001817) PMID: [24572233](https://pubmed.ncbi.nlm.nih.gov/24572233/).
18. Conte R et al. Manifesto of Computational Social Science. *Eur Phys J Spec Top*. 2012; 214: 325–346 doi: [10.1140/epjst/e2012-01697-8](https://doi.org/10.1140/epjst/e2012-01697-8)
19. Botta F, Moat HS, Preis T. Quantifying crowd size with mobile phone and Twitter data. *R Soc Open Sci*. 2015; 2: 150162 doi: [10.1098/rsos.150162](https://doi.org/10.1098/rsos.150162) PMID: [26064667](https://pubmed.ncbi.nlm.nih.gov/26064667/).
20. Barchiesi D, Moat HS, Alis C, Bishop S, Preis T. Quantifying international travel flows using Flickr. *PLOS ONE*. 2015; 10: e0128470 doi: [10.1371/journal.pone.0128470](https://doi.org/10.1371/journal.pone.0128470) PMID: [26147500](https://pubmed.ncbi.nlm.nih.gov/26147500/).
21. Alanyali M, Moat HS, Preis T. Quantifying the relationship between financial news and the stock market. *Sci Rep*. 2013; 3: 3578 doi: [10.1038/srep03578](https://doi.org/10.1038/srep03578) PMID: [24356666](https://pubmed.ncbi.nlm.nih.gov/24356666/).
22. Preis T, Moat HS, Bishop S, Treleaven S, Stanley HE. Quantifying the digital traces of hurricane Sandy on Flickr. *Sci Rep*. 2013; 3: 3141 doi: [10.1038/srep03141](https://doi.org/10.1038/srep03141) PMID: [24189490](https://pubmed.ncbi.nlm.nih.gov/24189490/).
23. Preis T, Moat HS. Adaptive nowcasting of influenza outbreaks using Google searches. *R Soc Open Sci*. 2014; 1: 140095 doi: [10.1098/rsos.140095](https://doi.org/10.1098/rsos.140095) PMID: [26064532](https://pubmed.ncbi.nlm.nih.gov/26064532/).
24. Preis T, Moat HS, Stanley HE, Bishop SR. Quantifying the Advantage of Looking Forward. *Sci Rep*. 2012; 2: 350 doi: [10.1038/srep00350](https://doi.org/10.1038/srep00350) PMID: [22482034](https://pubmed.ncbi.nlm.nih.gov/22482034/).
25. Preis T, Kenett DY, Stanley HE, Helbing D, Ben-Jacob E. Quantifying the behavior of stock correlations under market stress. *Sci Rep*. 2012; 2: 752 doi: [10.1038/srep00752](https://doi.org/10.1038/srep00752) PMID: [23082242](https://pubmed.ncbi.nlm.nih.gov/23082242/).
26. Preis T, Moat HS, Stanley HE. Quantifying trading behavior in financial markets using Google Trends. *Sci Rep*. 2013; 3: 1684 doi: [10.1038/srep01684](https://doi.org/10.1038/srep01684) PMID: [23619126](https://pubmed.ncbi.nlm.nih.gov/23619126/).
27. Moat HS, Curme C, Avakian A, Kenett DY, Stanley HE, Preis T. Quantifying Wikipedia usage patterns before stock market moves. *Sci Rep*. 2013; 3: 1801 doi: [10.1038/srep01801](https://doi.org/10.1038/srep01801)
28. Curme C, Preis T, Stanley HE, Moat HS. Quantifying the semantics of search behavior before stock market moves. *Proc Natl Acad Sci USA*. 2014; 111: 11600–11605 doi: [10.1073/pnas.1324054111](https://doi.org/10.1073/pnas.1324054111) PMID: [25071193](https://pubmed.ncbi.nlm.nih.gov/25071193/).
29. Münnix MC, Shimada T, Schäfer R, Leyvraz F, Seligman TH, Guhr T, Stanley HE. Identifying states of a financial market. *Sci Rep*. 2012; 2: 644 doi: [10.1038/srep00644](https://doi.org/10.1038/srep00644) PMID: [22966419](https://pubmed.ncbi.nlm.nih.gov/22966419/).
30. Mantegna R, Stanley HE. Scaling Behaviour in the Dynamics of an Economic Index. *Nature*. 1995; 376: 46–49 doi: [10.1038/376046a0](https://doi.org/10.1038/376046a0)
31. Gopikrishnan P, Plerou V, Amaral LAN, Meyer M, Stanley HE. Scaling of the distributions of fluctuations of financial market indices. *Phys Rev E*. 1999; 60: 5305–5316 doi: [10.1103/PhysRevE.60.5305](https://doi.org/10.1103/PhysRevE.60.5305)
32. Gabaix X. Power laws in economics and finance. *Annu Rev Econom*. 2009; 1: 255–294 doi: [10.1146/annurev.economics.050708.142940](https://doi.org/10.1146/annurev.economics.050708.142940)
33. Lux T, Marchesi M. Scaling and criticality in a stochastic multi-agent model of a financial market. *Nature*. 1999; 397: 498–500 doi: [10.1038/17290](https://doi.org/10.1038/17290)
34. Gabaix X, Gopikrishnan P, Plerou V, Stanley HE. A theory of power-law distributions in financial market fluctuations. *Nature*. 2003; 423: 267–270 doi: [10.1038/nature01624](https://doi.org/10.1038/nature01624) PMID: [12748636](https://pubmed.ncbi.nlm.nih.gov/12748636/).
35. Plerou V, Gopikrishnan P, Stanley HE. Two-phase behavior of financial markets. *Nature*. 2003; 421: 130 doi: [10.1038/421130a](https://doi.org/10.1038/421130a) PMID: [12520293](https://pubmed.ncbi.nlm.nih.gov/12520293/).
36. Podobnik B, Horvatic D, Petersen AM, Stanley HE. Cross-correlations between volume change and price change. *Proc Natl Acad Sci USA*. 2009; 106: 22079–22084 doi: [10.1073/pnas.0911983106](https://doi.org/10.1073/pnas.0911983106) PMID: [20018772](https://pubmed.ncbi.nlm.nih.gov/20018772/).
37. Gu GF, Chen W, Zhou W-X. Empirical distributions of Chinese stock returns at different microscopic timescales. *Physica A*. 2008; 387: 495–502 doi: [10.1016/j.physa.2007.10.012](https://doi.org/10.1016/j.physa.2007.10.012)
38. Makowiec D, Gnacinski P. Fluctuations of WIG—the index of Warsaw stock exchange preliminary studies. *Acta Phys*. 2001; 32: 1487–1500.
39. Bertram WK. An empirical investigation of Australian stock exchange data. *Physica A*. 2004; 341: 533–546 doi: [10.1016/j.physa.2004.04.132](https://doi.org/10.1016/j.physa.2004.04.132)
40. Coronel-Brizio HF, Hernandez-Montoya AR. On fitting the Pareto-Levy distribution to stock market index data—Selecting a suitable cutoff value. *Physica A*. 2005; 354: 437–449 doi: [10.1016/j.physa.2005.03.001](https://doi.org/10.1016/j.physa.2005.03.001)

41. Plerou V, Stanley HE. Tests of scaling and universality of the distributions of trade size and share volume: Evidence from three distinct markets. *Phys Rev E*. 2007; 76: 046109 doi: [10.1103/PhysRevE.76.046109](https://doi.org/10.1103/PhysRevE.76.046109)
42. Mu GH, Zhou W-X. Tests of nonuniversality of the stock return distributions in an emerging market. *Phys Rev E*. 2010; 82: 066103 doi: [10.1103/PhysRevE.82.066103](https://doi.org/10.1103/PhysRevE.82.066103)
43. Axtell RL. Zipf distribution of US firm sizes. *Science*. 2001; 293: 1818–1820 doi: [10.1126/science.1062081](https://doi.org/10.1126/science.1062081) PMID: [11546870](https://pubmed.ncbi.nlm.nih.gov/11546870/).
44. Sornette D, Woodard R, Zhou WX. The 2006–2008 oil bubble: evidence of speculation, and prediction. *Physica A*. 2009; 388: 1571–1576 doi: [10.1016/j.physa.2009.01.011](https://doi.org/10.1016/j.physa.2009.01.011)
45. Clauset A, Shalizi CR, Newman MEJ. Power-law distributions in empirical data. *SIAM Rev*. 2009; 51 (4): 661–703 doi: [10.1137/070710111](https://doi.org/10.1137/070710111)
46. Yan C, Zhang JW, Zhang Y, Tang YN. Power-law properties of Chinese stock market. *Physica A*. 2005; 353: 425–432 doi: [10.1016/j.physa.2005.02.010](https://doi.org/10.1016/j.physa.2005.02.010)
47. Zhang JW, Zhang Y, Kleinert H. Power tails of index distributions in Chinese stock market. *Physica A*. 2007; 377: 166–172 doi: [10.1016/j.physa.2006.11.012](https://doi.org/10.1016/j.physa.2006.11.012)
48. Matia K, Pal M, Salunkay H, Stanley HE. Scale-dependent price fluctuations for the Indian stock market. *Europhys Lett*. 2004; 66: 909–914 doi: [10.1209/epl/i2003-10267-y](https://doi.org/10.1209/epl/i2003-10267-y)